DOCUMENT RESUME

ED 133 698 CS 003 139

AUTHOR Harber, Jean R.
TITLE Reading Tests and The Dialect Speaker.
PUB DATE 76
NOTE 22p.; Report prepared at University of Maryland

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Culture Free Tests; Elementary Secondary Education; *Evaluation Criteria; *Nonstandard Dialects; *Reading Tests; *Standardized Tests; Standard Spoken Usage; *Test Bias; *Testing Problems; Test Reviews; Test Selection

ABSTRACT

Characteristics which should be considered in reviewing any standardized reading test include validity, reliability, standardization sample, areas assessed by the test, type of response required by the child, individual versus group administration, time needed for administration, availability of equivalent forms, scoring options available, reviewers' comments, information given concerning interpretation of results and/or instructional suggestions, and groups for whom the test is or is not appropriate. These characteristics are presented in table form as they apply to 14 reading tests, including the Gates-MacGinitie Reading Tests (Readiness), the Metropolitan Readiness Tests, the Woodcock Reading Mastery Tests, and the Stanford Achievement Tests. Empirical evidence suggests that there is a need to develop either dialect forms of reading tests or alternative scoring procedures for dialect speakers. (KS)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED133698

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Reading Tests

and

The Dialect Speaker

Jean R. Harber

CS003139

Reading tests have been used for many years to assess pupil achievement, determine pupil readiness and to identify specific strengths and weaknesses. Teachers and other school personnel often select one or more published tests from those available to them and administer these tests to children in an attempt to assess the children's achievement, readiness or specific strengths and weaknesses. A test should be carefully reviewed before a decision is made to use or reject it for a given purpose. The criteria which should be considered in reviewing any reading test includes: validity, reliability, standardization sample, areas assessed by the test, type of response required of child, individual versus group administration, time needed to administer test, availability of equivalent forms, scoring options available, reviewers' comments, information given concerning interpretation of results and/or instructional suggestions and groups for whom test is or is not appropriate. Each is discussed briefly below:

Validity answers the question, "To what extent does the test measure what it purports to measure?" Validity can be measured in several ways. Content validity refers to the extent to which the test taps knowledge of the curricular content and cognitive processes. Content validation studies are commonly carried out when achievement tests are constructed. Criterion-related (or predictive) validity tells how well a test measures future performance on some criterion. It is particularly important in readiness tests. Construct validity tells the degree to which certain psychological traits or constructs are actually represented by test performance. Face validity refers to whether or not on first impression, the test appears to measure the intended content.

Reliability refers to the accuracy of a measuring instrument. It answers the question, "If we measure the same set of objects again and again with the same instrument, will we get the same or similar results?" Reliability, too, can be measured in several ways. Determining test-retest reliability involves administering the same test twice to see if individual scores change from one time to the next. Determining parallel-form reliability involves administering two forms of a test which are considered to be equivalent to determine whether scores change from one test to the other. Split-half reliability is determined through one test administration. The test is divided into equal parts (e.g., odd-even) before it is administered and the two parts are treated as if they were separate tests. The higher the coefficient, the more reliable the instrument. The standard error of measurement of a test indicates how chance errors may cause variations in the scores which might be obtained by an individual if the same test were administered numerous times. It is desirable to use a test with a relatively small standard error.

To standardize a test, publishers administer it to a large group of students selected to be representative of the population at the grade level(s) for which the test is intended. This group is typically called the standardization sample. In some cases a major effort is made to see that the sample is representative of all the students in the country. In other cases a much less representative sample is taken. It is not the size of the sample that is of primary importance, but rather the sample's representativeness of the group(s) the test is intended for.

One needs to be aware of the areas assessed by a given test in order to be able to match a test to the needs of the prospective examiner. Clearly, if a teacher wanted to measure silent reading comprehension,

he should not use an oral reading test. One must also consider the type of response required of the student. Does the child have difficulty filling in small circles or responding verbally which will negatively influence the score he obtains? The setting in which the test is administered (individually or in a group) and the time it takes to administer the test should also be considered. If the teacher wants to give the test as a pretest and later as a posttest, a test which has two or more equivalent forms would be advantageous. The teacher also needs to consider the kinds of scores he wants to obtain and whether a given test yields the results in the specified format(s). When available, reviewers' comments should be considered as well.

Test manuals vary in the amount of information given concerning the interpretation of results and/or instructional suggestions. Clearly, it is helpful to be provided with this information. Finally, one must be concerned with whether a test is appropriate for the various dialectal groups in our schools today.

Numerous reading tests were reviewed according to these criteria and are summarized in Table I.

INSERT TABLE I ABOUT HERE

As shown in Table I, none of the tests reviewed take into account dialect differences. Research has shown that certain tests are linguistically and culturally biased (Hutchinson, 1972; Roberts, 1970). There is some evidence in the literature (Hunt, 1975) which indicates that Black English-speaking subjects scored significantly higher on an oral reading test presented in standard English when "errors"

attributable to dialect were not counted as errors than when tests were socred according to the directions given in the test manual. Harber (1975) found that Black English-speaking subjects scored significantly higher on oral reading passages presented in black English standard orthography than on equivalent oral reading passages presented in standard English. Thus, there is empirical evidence which suggests that there is a need for developing either dialect forms of reading tests or alternative scoring procedures for dialect speakers. Not to provide such tests or scoring procedures could lead in inaccurate and misleading reading evaluations and inappropriate classification and placement of dialect speakers.

Table I

Characteristics of Tests Reviewed

Name	Type of Test	Standardization Sample	Reliability Data Given In Terms of	Validity Data Given In Terms of	Individual Or Group Administration	Type of Response Required	Time Taken to Administer	Equivalent Forms Available	Grade Levels Appropriate For
Gates-MacGinitie Reading Tests: Readiness	R	LNS	AF(.83-.86) SH(.91-.94)	Not reported	G	Gr	120min.	Yes-2	End of K & beginning First Grade
Metropolitan Readiness Tests	R	LNS	AF(.91) SH(.90-.95)	Content Construct Predictive	G	Gr	60min.	Yes-2	K & beginning First grade
Murphy-Durrell Reading Readiness Analysis	R	LNS	SH(.88-.98) SE_m(1.5-3.5)	Content Predictive Concurrent	G	Gr	60min.	No	Beginning First Grade
Durrell Listening Reading Series: Primary Level	S	LNS	SH(.96) SE_m(4.2-6.7)	Content Construct Concurrent	G	Gr	70min.	No	Grades 1 through middle of Grade 3
Gates-MacGinitie Reading Tests Primary Level	S	LNS	AF(.83-.86) SH(.91-.94)	Not reported	G	Gr	40min.	Yes-1	Grades 1-3
Gray Oral Reading Tests	S	Small	SE_m(3.04-4.37)	Concurrent	I	V	10-15 min.	Yes-4	Grade 1-Adults

Table I

Characteristics of Tests Reviewed

Areas Assessed	Demonstration Items Provided	Scores Available	Detailed Discussion on Interpretation of results for instructional suggestions included in manual	Takes into account dialect differences	Criticisms stated by Reviewers
Listening comprehension Auditory discrimination Visual discrimination Following directions Letter recognition Visual-motor coordination Auditory blending Word recognition	Yes	GE NNS PR RS	No	No	No clues given to child to help him identify the line which the examiner is describing (Berg, 1972).
Word meaning Listening Matching Alphabet Numbers Copying	Yes	LR PR RS S	Yes	No	Teachers do not gain very much educationally useful information because the test predicts future achievement on the basis of past achievement (Singer, 1972 a).
Phonemes Leter names Learning rate	Yes	PR QR RS S	Yes	No	The basic assumption of the learning rate test may be only partially valid (Singer, 1972 b).
Vocabulary listening Sentence listening Vocabulary reading Sentence reading	Yes	AE GE PR RS S	Yes	No	Construct and concurrent validity are weak (Spache, 1972).
Vocabulary Comprehension	Yes	GS NSS PR RS	No	No	Vocabulary section requires child make very graphic & phonemic contrasts between words with minimal contrasts. Reading selections on the comprehension section are too short/ (Burke, 1972).
Oral reading Literal oral reading comprehension	No	GE RS SE	Yes	No	"Tentative" norms are based on a small sample (Harris, 1965; Lohnes, 1965).

Dialect Speaker 7

Table I (contd.)

Characteristics of Tests Reviewed

Name	Type of Test	Standardization Sample	Reliability Data Given In Terms of	Validity Data Given In Terms of	Individual Or Groups Administration	Type of Response Required	Time Taken To Administer	Equivalent Forms Available	Grade Levels Appropriate For
Durrell Analysis of Reading Difficulty:Primary Level	D	LNS	Not reported	Not reported	I	V	30-90min.	No	Grade 1 through middle of grade 3
Gates-McKillop Reading Diagnostic Tests	D	Not-reported	Not reported	Not reported	I	Gr & V	30-45min.	Yes-2	Elementary Grades & beyond
Stanford Diagnostic Reading Test:Level I	D	LNS	SH(.85-.96)	Content Construct Concurrent	G	Gr	140min.	Yes-2	Middle of Grade 2 to middle of Grade 4

Table I (contd.)

Characteristics of Tests Reviewed

Areas Assessed	Demonstration Items Provided	Scores Available	Detailed Discussion on Interpretation of results for instructional suggestions included in manual	Takes into account dialect differences	Criticisms stated by Reviewers
Oral reading Silent reading Listening comprehension Word Recognition and analysis Visual memory of word forms Auditory analysis of word elements Spelling Handwriting	Yes	GE RS	Yes	No	No reliability or validity information given in manual (Robinson, 1953).
Oral Reading Words-Flash presentation Words-untimed presentation Phrases-flash presentation Knowledge of word parts Recognizing visual forms of sounds Auditory blending Spelling Oral vocabulary Syllabication Auditory discrimination	No	AE GE R RS	Yes	No	
Reading comprehension Vocabulary Auditory discrimination Syllabication Beginning & ending sounds Blends Sound discrimination	Yes	GE PR RS S	Yes	No	

Table I (contd.)
Characteristics of Tests Reviewed

Name	Type of Test	Standardization Sample	Reliability Data Given In Terms of	Validity Data Given In Terms of	Individual Or Group Administration	Type of Response Required	Time Taken To Administer	Equivalent Forms Available	Grade Levels Appropriate For
Woodcock Reading Mastery Tests	CR	LNS	AF(.83-.97) SH(.98-.99) SE_m(1.1-3.1)	Content Multimethod-Multitrait Matrix Predictive	I	V	20-30min.	Yes-2	K-12
Peabody Individual Achievement Test	SC	LNS	SE_m(3.06-6.51) TR(.64-.89) (reading sub-tests)	Content Concurrent	I	P V	30-40min.	No	Preschool through High School
California Achievement Tests	S	LNS	High reliability	Not reported	G	Gr	240min.	Yes-2	Middle of Grade 1 through Grade 2
Metropolitan Achievement Tests	S	LNS	SH(.87-.93) SE_m(2.3-2.7) (reading subtests)	Content	G	Gr	100min.	yes-2	Second half of Grade 1
Stanford Achievement Tests	S	LNS	SH(high enough for diagnosing individual problems)	Little solid evidence available	G	Gr	120-185 min.	Yes-3	Second half of Grade 1 and first half of Grade 2

Table I (contd.)

Characteristics of Tests Reviewed

Areas Assessed	Demonstration Items Provided	Scores Available	Detailed Discussion on Interpretation of results for instructional suggestions included in manual	Takes into account dialect differences	Criticisms stated by Reviewers
Letter identification Word identification Word attack Word comprehension Passage comprehension	Yes	AE AI GE MS NNS PR RS	Yes	No	
Reading recognition Reading comprehension Math General information	Yes	AE GE PR RS NSS	No	No	
Reading vocabulary Reading comprehension Math Language	Yes	GE PR RS S	Yes	No	
Word knowledge Word discrimination Reading Arithmetic	Yes	GE NSS PR R RS S	Yes	No	
Word meaning Paragraph meaning Science & social studies concepts Spelling Word study skills Language Arithmetic	Yes	GE PR RS S	Yes	No	Very little solid evidence concerning the validity of the test is available (Traxler, 1972).

Table I (contd.)

Code:

Type of Test

- CR - Criterion-referenced
- D - Diagnostic
- R - Readiness
- S - Survey
- SC - Screening

Standardization Sample

- LNS - Large national sample

Reliability Data given in terms of

- AF - Alternate form
- SH - Split half
- SE - Standard error of measurement
- TR^m - Test-retest

Individual or group administration

- G - Group
- I - Individual

Type of response required

- Gr - Graphic
- P - Pointing
- V - Vocal

Scores available

- AE - Age equivalent scores
- AI - Achievement index
- GE - Grade equivalent scores
- LR - Letter ratings
- MS - Mastery scores
- NSS - Normalized standard scores
- PR - Percentile ranks
- QR - Quartile ratings
- R - Ratings
- RS - Raw scores
- S - Stanines
- SE - Standard error

References

Berg, P. C. Gates-MacGinitie Reading Tests: Readiness Skills. In O.K. Buros (Ed.), The Seventh Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1972.

Burke, C. L. Gates-MacGinitie Reading Tests. In O.K. Buros (Ed.), The Seventh Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1972.

Harber, J. R. The effects of abstract reasoning ability, degree of bidialectism, and grade level on the performance of black, inner-city children on equivalent reading tasks presented in standard English, Black English standard orthography, and Black English nonstandard orthography. Unpublished doctoral dissertation, Temple University, 1975.

Harris, A.J. Gray Oral Reading Tests. In O.K. Buros (Ed.), The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1965.

Hunt, B. C. Black dialect and third and fourth graders' performance on the Gray Oral Reading Test. Reading Research Quarterly, 1975, 10, 103-123.

Hutchinson, J.O. Reading tests and nonstandard language. Reading Teacher, 1972, 25, 430-437.

Lohnes, P.R. Gray Oral Reading Test. In O.K. Buros (Ed.), The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1965.

Roberts, E. An evaluation of standardized tests as tools for the measurement of language development. Unpublished paper, Language Research Foundation, Cambridge, 1970.

Robinson, H.M. Durrell Analysis of Reading Difficulty. In O.K. Buros (Ed.), The Fourth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1953.

Singer, H. Metropolitan Readiness Tests. In O.K. Buros (Ed.), The Seventh Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1972. (a)

Singer, H. Murphy-Durrell Reading Readiness Analysis. In O.K. Buros (Ed.), The Seventh Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1972. (b)

Spache, G.D. Durrell Listening-Reading Senses. In O.K. Buros (Ed.), The Seventh Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1972.

Traxler, A.E. Stanford Achievement TEst: Reading Tests. In O.K. Buros (Ed.), The Seventh Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1972.

Test References

Dunn, L. M. & Markwardt, F. C. Peabody Individual Achievement Test, Circle Pines, Minnesota: American Guidance Service, 1970.

Durost, W. N., Bixler, H. H. Hilbreth, G. H. Lund, K. W. & Wrightstone, J. W. Metropolitan Achievement Tests: New York: Harcourt, Brace, & World, 1959.

Durrell, D. D. Durrell Analysis of Reading Difficulty: New York: Harcourt, Brace & World, 1955.

Durrell, D. D. & Hayes, M. T. Durrell Listening-Reading Series. New York: Harcourt, Brace, & World, 1969.

Gates, A. I., & MacGinitie, W. H. Gates-MacGinitie Reading Tests. New York: Teachers College, Columbia University: 1965.

Gates, A. I. & McKillop, A. S. Gates-McKillop Reading Diagnostic Tests. New York: Teachers College Columbia University, 1962.

Gray, W. S. Gray Oral Reading Tests. Indianopolis: Bobbs-Merrill, 1963.

Hildreth, G. G., Griffiths, N. L. & McGauvran, M.E. Metropolitan Readiness Tests. New York: Harcourt, Brace, & World, 1976.

Karlsen, B., Madden, R. & Gardner, E. F. Stanford Diagnostic Reading Test. New York: Harcourt, Brace, & World, 1966.

Kelley, T. L., Madden, R. Gardner, E.F. & Rudman, H. C., Stanford Achievement Test. New York: Harcourt, Brace, & World, 1963.

Tiegs, E. C. & Clark, W. W. California Achievement Tests. Monterey, California: ETB McGraw-Hill, 1970.

Woodcock, R. W., Woodcock Reading Mastery Tests. Circle Pines, Minesota: American Guidance Service, 1973.

Murphy, H. A. & Durrell, D.D. Murphy-Durrell Reading Readiness Analysis. New York: Harcourt, Brace & World, 1965.